

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 03:30:35 ; Search time 1713.89 Seconds  
(without alignments)  
3709.145 Million cell updates/sec

Title: US-09-525-361A-23

Perfect score: 471

Sequence: 1 cttgaagcatttttctgt.....aactatgacgagctaacat 471

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_estl:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_gss:\*
- 13: em\_gss\_hum:\*
- 14: em\_gss\_inv:\*
- 15: em\_gss\_pln:\*
- 16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	431.8	91.7	436	10	W72837
2	403.2	85.6	432	10	W72838
3	330.8	83.0	420	9	AW070916
4	388.6	82.5	423	10	BF742809
5	379	80.5	402	9	AI139456
6	345.2	73.3	464	10	BG202312
7	331	70.3	376	10	BG218084
8	316.8	67.3	345	9	BE092421
9	309.8	65.8	391	10	BE185269
10	308.8	65.6	403	10	R75793
11	301.4	64.0	314	12	AZ694036
12	299.2	63.5	343	10	BG202313
13	294	62.4	313	9	AW176044
14	289.4	61.4	294	9	AA340069
15	281	59.7	316	9	BE186013
16	281	59.7	316	10	BE926938
17	255.8	54.3	443	9	AI127172

18	247.6	52.6	308	9	AI905687
19	217.8	46.2	251	10	BE197186
20	202.8	43.1	221	9	BE074512
21	197.8	42.0	462	9	AW294149
22	197	41.8	224	9	AW291950
23	196.6	41.7	230	10	BE815819
24	187.8	39.9	248	10	BE815824
25	159	33.8	193	9	AI905624
26	154	32.7	224	9	AI905837
27	151	32.1	211	10	BE181376
28	150	31.8	212	10	BE192597
29	149.4	31.7	178	9	AI905623
30	140.4	29.8	211	10	BG207535
31	133	28.2	136	9	AI905633
32	102.8	21.8	224	10	BE815853
33	85	18.0	498	10	BE1279448
34	83	17.6	536	10	BI339750
35	82.6	17.5	572	10	BF080542
36	82.4	17.5	490	10	BI279279
37	82.4	17.5	492	10	BI279431
38	82.4	17.5	499	10	BI279473
39	81.8	17.4	486	10	BI282803
40	80.8	17.2	481	10	BI279272
41	80.8	17.2	490	10	BI279159
42	80.8	17.2	493	10	BI279513
43	80.8	17.2	493	10	BI279518
44	80.8	17.2	495	10	BI279214
45	80.8	17.2	496	10	BI279121

## ALIGNMENTS

RESULT 1

W72837

LOCUS

DEFINITION

W72837

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

W72837 436 bp mRNA linear EST 16-OCT-1996  
zds57g12.r1 Soares\_fetal\_heart\_NBHH19W Homo sapiens cDNA clone  
IMAGE:344806 5' similar to contains element MER40 repetitive  
element ;, mRNA sequence.

W72837 GI:1382813

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 436)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman

, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,

Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston

, R., Williamson, A., Wohldmann, P. and Wilson, R.

The Washu-Merck EST Project

Unpublished (1995)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 508 Std Error: 0.00

Seq primer: mob.REGA+ET.

Location/Qualifiers

1..436

/organism="Homo sapiens"

/db\_xref="GDB:1270181"

/db\_xref="taxon:9606"

/clone\_lib="IMAGE:344806"

/sex="unknown"

/dev\_stage="19 weeks"

/lab\_host="DH10B (ampicillin resistant)"



```
RESULT 1
AW070916/c
LOCUS
DEFINITION
  AW070916.1 NC1_CGAP_Br18 Homo sapiens cDNA clone IMAGE:2568385 3'
  similar to contains element TARI repetitive element ;, mRNA
  sequence.
ACCESSION
  AW070916
VERSION
  AW070916.1 GI:6025914
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 420)
REFERENCE
  NC1_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgaps@mail.nih.gov
  Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.
  Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life
  Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
  I.M.A.G.E. Consortium DNA Sequencing by: Washington University
  Genome Sequencing Center
  Clone distribution: NC1_CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www.bio.llnl.gov/bbrp/image/image.html
  Insert length: 545 Std Error: 0.00
  Seq primer: -40UP from Gbco.
  Location/Qualifiers
    1..420
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:2568385"
      /clone_lib="NC1_CGAP_Br18"
      /tissue_type="four pooled high-grade tumors, including two
      primary tumors and two metastatic to ovary"
      /lab_host="PH10B"
      /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: Salt;
      Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
      Library constructed by Life Technologies."
  BASE COUNT
    119 a 76 c 116 g 109 t
  ORIGIN
    57 tagcagctcctggtactcttggaggtttccatcttctgtctctgcccagaatccgacaa 116
    |||||||
    Db 420 TACCAGTCTCGTACTCTTGGGAGTTTCCATCTTCTGTCTCTGCCCGAATCCGACAA 361
    |||||||
    Qy 117 cagctgctccagctgacagctatccagctactggtctgctgctgatgatgaagccctgatg 176
    |||||||
    Db 360 CAGCTGCTCCAGCTGACAGCTATCCAGCTACTGGTCTCTGCTGATGATGAAGCCCTGATG 301
    |||||||
    Qy 177 ctgaaacacactgctgtgcaaacactgacacactgctgctctcctaccctgcaacacgcg 236
    |||||||
    Db 300 CTGAACACCACTGCTGTGCAACCACTGCCACCACTGCTGCTCTCTCTCTCTCTCTCTCT 241.
    |||||||
    Qy 237 ctgcttctaccctgctgtgaaagacatccagttttaccacaaagggtggggtctcc 296
    |||||||
    Db 240 CTGCTTCTACCACTGCTGCTAAAGACATTCAGTTTACCCAAATGGGTGGGGATCTCC 181
    |||||||
    Qy 297 cgaatggtagtggtgtccctgagatggaatcagcttgagttctcttgcgaattgggtccaca 356
    |||||||
    Db 180 GCAATGGTAGAGTGTCTCCCTGAGATGGAATCAGCTTGAGTTAGTCTTGCATTT-GGTACA 122
    |||||||
    Qy 357 actattcagctctctgtgatttcacaaactactactacccttccctcagcatatcccccttta 416
    |||||||
```

```
Db 121 ACTATTTCATGCTTCCTGCTGATTTCATCCAACTACTTACCTTGCCTACGATATCCCTTTTA 62
Qy 417 tctctaatacagtttatttcttcttcaataaaaaataactatgagcagctaaaca 470
|||||
Db 61 TCTCTAATCAGTTTATTCTTCTTCAATAAAAAATAACTATGAGCAACAAAAA 8
|||||

RESULT 4
BF742809
LOCUS
DEFINITION
  IL2-BT0807-041000-176-G03 BT0807 Homo sapiens cDNA, mRNA sequence.
ACCESSION
  BF742809
VERSION
  BF742809.1 GI:12069485
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 423)
REFERENCE
  Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
  Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
  Goldman,G.H., Carvalho,A.F., Matsukuma,A., Balia,G.S., Simpson,D.H.,
  Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
  ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
  Simpson,A.J.
  Shotgun sequencing of the human transcriptome with ORF expressed
  sequence tags
  Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  20202663
  Contact: Simpson A.J.G.
  Laboratory of Cancer Genetics
  Ludwig Institute for Cancer Research
  Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
  Brazil
  Tel: +55-11-2704922
  Fax: +55-11-2707001
  Email: asimpson@ludwig.org.br
  This sequence was derived from the FAPESP/LICR Human Cancer Genome
  Project. This entry can be seen in the following URL
  (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=IL2&t2=IL2-BT0807-
  041000-176-G03&t3=2000-10-04&t4=1)
  Seq primer: puc 18 forward
  High quality sequence stop: 421.
  Location/Qualifiers
    1..423
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone_lib="BT0807"
      /dev_stage="Adult"
      /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
      SmaI; A mini-library was made by cloning products derived
      from ORESTES PCR (U.S. Letters Patent application No. 196
      ,716 - Ludwig Institute for Cancer Research) profiles
      into the pUC 18 vector. Reverse transcription of tissue
      mRNA and cDNA amplification were performed under low
      stringency conditions."
  BASE COUNT
    89 a 123 c 87 g 124 t
  ORIGIN
    Query Match 82.5%; Score 388.6; DB 10; Length 423;
    Best Local Similarity 98.8%; Pred. No.1e-89;
    Matches 402; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
    Qy 1 ctttgaagcattttgtctgtgctccctgatcttcagctcaccaccatgaagttcttagc 60
    |||||||
    Db 15 CTTTGAAGCATTTTGTGCTGCTGCTCCCTGATCTTCAGTCCACCACCATGAAGTTCTTAGC 74
    |||||||
    Qy 61 agtctcgttactctctggaggttccatcttctgtgctctcccaagaatccgcacaacgc 120
    |||||||
    Db 75 AGTCTGTGTACTCTTGGGAGTTTCCATCTTTCTGTGCTCTCTGCCAGAAATCCGACAACGC 134
    |||||||
    Qy 121 tgcctcagctgacagctatccagctactggtccctgctgatgatgaagccctctgatgctga 180
    |||||||
```



Query Match	73.3%	Score	345.2	DB 10:	Length	464			
Best Local Similarity	97.6%	Pred. No.	1.6e-78:						
Matches	361	Conservative	0	Mismatches	8	Indels	1	Gaps	1
QY	101	qccagaatccgacacagctgctccagctgacacg	tatccagctactcggctcgtcgtgat	160					
DB									
DB	89	GCACAGAAATCCGACACAGCTGCTCCAGCTC	CACAGTATCCAGGTACTGGTCTGCTGAT	148					
QY	161	gataagccctcgatcgtgaaacccactgctcgl	cgaacaactcgacacactgctgctcct	220					
DB	149	GATGAAGCCCTGATGCTGAAACCACTGCTG	CTGCTGCAACCACTGCGACCACTGCTGCTCT	208					
QY	221	accactgcaacccacccgctgctctaccactg	ctcgtgaaagacattccagttttacccaaa	280					
DB	209	ACCACTGCAACCAACCGGTGCTTCTACCACT	GCTCTGTAAGAGACATTTCCAGTTTTACCCAAA	268					
QY	281	tyggctggggatctcccgaatgggagagtg	tgtcccctgagatggaaatcagctgaagctt	340					
DB	269	TGGGTTGGGGATCTCCCGAATGGTAGACT	GTGTCTCCCTGAGATGGAATCAGCTTGAGTCTT	328					
QY	341	ctgcaatggggctcaactattcagcttcctc	gtgatttcacaaactacttaccctggcc	400					
DB	329	CTGCAATTT-GGTCAACAATAATCATGCT	TCTGCTGTGATTTCTCAACACTACTTACTCTTGGC	387					
QY	401	taqatatacccccttatctctaatacagttat	ttctttcttcaataaaaaataaactatgag	460					
DB	388	TACGATATCCCTTTATCTCTAATCAGTTT	ATTATTTCTTTCAATAAAAAATAACTATGAG	447					
QY	461	cgagctaaca	470						
DB	448	CAACAAAAAA	457						

RESULT 7  
 BG218084  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 376)  
 Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
 Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,  
 Lerner,L., Costanzo,D., McGilligott,K., Booser,S., Mays,R., Smith,  
 F., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,  
 J., Danzig,J. and Ducar,M.  
 Creation of genome-wide protein expression libraries using random  
 activation of gene expression  
 Nat. Biotechnol. 19 (5), 440-445 (2001)  
 21227151  
 Contact: Scott J. Cain  
 TITLE  
 JOURNAL  
 MEDLINE  
 COMMENT

**FEATURES**  
 source  
 Athersys, Inc.  
 3201 Carnegie Ave., Cleveland, OH 44115, USA  
 Tel: 216 431 9900  
 Fax: 216 361 9596  
 Email: scaine@atersys.com  
 High quality sequence stop: 376.  
 Location/Qualifiers  
 1..376  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Athersys RAGE Library"  
 /cell\_line="HT1080"  
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily

FEATURES  
SOURCE

High quality sequence start: 9  
High quality sequence stop: 345.

Location/Qualifiers

source  
1..345  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="BT0738"  
/dev\_stage="Adult"  
/note="Organ: breast; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 92 a 77 c 103 g 73 t  
ORIGIN

Query Match 67.3%; Score 316.8; DB 9; Length 345;  
Best Local Similarity 97.6%; Pred. No. 3.2e-71;  
Matches 332; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 37 ggccacacacatgaattcttaacagctcctggtactcttgaggagtttccattcttctgtg 96  
|||||  
Db 345 ggccacacacatgaattcttaacagctcctggtactcttgaggagtttccattcttctgtg 286  
|||||  
QY 97 ctctgcccagaatccgacacacagctcctccagctgacacgtatccagctactggtcctgc 156  
|||||  
Db 285 CTCTGCCCAATCCGACACACAGCTCTCCAGCTGACACGTATCCAGCTACTGTCCTGC 226  
|||||  
QY 157 tcatgatgaagccctgatgctgaaccactgctgtgcaacactgcgaccactgctgc 216  
|||||  
Db 225 TGATGATGAAGCCCTGATGCTGAAACCACTGCTGTCGAACCACTCCGACCACTGCTGC 166  
|||||  
QY 217 tcttaccactgaaccacacgctcttaccactgctgtaagacattccagctttacc 276  
|||||  
Db 165 TCCTACCACCTGCAACCAACCCGCTGCTTACCACTGCTGTAAGACATTCACGTTTACC 106  
|||||  
QY 277 caaatgggtgggagatcccccgaatggtagagtgctccctcctgagatggaaatcagcttgag 336  
|||||  
Db 105 CAAATGGTTGGGATCTCCGAATGCTAGAGTGTCT-CCTGAGATGGAATCAGCTTGAG 47  
|||||  
QY 337 tcttctgcaattgggtgcacaaactatcatgcttccctgtga 376  
|||||  
Db 46 TCTTCTGCTAGTGGTGCACAACTATTTCATGCTTCCCTCGA 7  
|||||

RESULT 9

LOCUS BG185269 391 bp mRNA linear EST 21-APR-2001  
DEFINITION R574208 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.

ACCESSION BG185269

VERSION BG185269.1 GI:13706956

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo;

1 (bases 1 to 391)

Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,

Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.,

Lerner, L., Costanzo, D., McElligott, K., Boozer, S., Mays, R., Smith

, E., Veloso, N., Kika, A., Hess, J., Cothren, K., Lo, K., Offenbacher

, J., Danzig, J. and Ducar, M.

Creation of genome-wide protein expression libraries using random

activation of gene expression

Nat. Biotechnol. 19 (5), 440-445 (2001)

21227151

Contact: Scott J. Cain

Athersys, Inc.

3201 Carnegie Ave., Cleveland, OH 44115, USA

Tel: 216 431 9900

Fax: 216 361 9596

Email: scain@athersys.com

High quality sequence stop: 391.

Location/Qualifiers

source  
1..391  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Athersys RAGE Library"  
/cell\_line="HT1080"  
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

BASE COUNT 105 a 103 c 72 g 101 t  
ORIGIN

Query Match 65.8%; Score 309.8; DB 10; Length 391;  
Best Local Similarity 97.1%; Pred. No. 2.1e-69;  
Matches 336; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 125 ccagctgacacgtatccagctactgctcctgctgatgagagccctgagctgaaacc 184  
|||||  
Db 41 CCAGCTGNAACGTATCCAGCTACTGCTCCTGCTGATGATGAAGCCCTCGA-GCTCAAAACC 99  
|||||  
QY 185 actgctgctgcacaaactgcaccactgctcctcctaccactgcaaccaccgctctct 244  
|||||  
Db 100 ACTGCTGCTGCAACCACTCCGACCACTGCTGCTCTTACCACCTGCAACCACTGCTGCT 159  
|||||  
QY 245 accactgctgtaagacacattccagctttaccacaaatgggtgggagatcccccgaatggt 304  
|||||  
Db 160 ACCACTGCTCGTAAACACATTCAGTTTTTACCAGTTGGTGGGATCTCCGAAATGCT 219  
|||||  
QY 305 agagtgctccctgagatgagatcagctgagcttcttgcgaattgggtgcacaaacttca 364  
|||||  
Db 220 AGAGTGCTGCCCTGAGATGGAATCAGCTTGAGTCTTCTGCAATT-GGTCAACACTATTCA 278  
|||||  
QY 365 tcttctcctgatttcatcccaactacttaccctgctcagcatatccccctttatctctaat 424  
|||||  
Db 279 TGCTTCTCTGATTTCATCCAACTACTTACTTACCTGCTACGATATCCCCCTTTATCTCTAAT 338  
|||||  
QY 425 cagtttattttctttccaaataaaaaataactatgagcagagctaaaca 470  
|||||  
Db 339 CAGTTTATTTCTTTCAAAATAAAAAATAACTATGAGCAACAAAAA 384  
|||||

RESULT 10

R57593

LOCUS

DEFINITION

IMAGE:158921 5' similar to contains LTR3 repetitive element ;, mRNA

sequence.

ACCESSION R57593

VERSION R57593.1 GI:850475

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 403)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman

, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,

Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston

, R., Williamson, A., Wohldmann, P. and Wilson, R.

The WashU-Merck EST Project

Unpublished (1995)

JOURNAL

COMMENT

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu  
InsertL Size: 516  
High quality sequence stops: 307  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 516 Std Error: 0.00  
Seq primer: M13RP1  
High quality sequence stop: 307.  
Location/Qualifiers  
1. 403  
/organism="Homo sapiens"  
/db\_xref="GDB:572980"  
/db\_xref="taxon:9606"  
/clone="IMAGE:158921"  
/clone\_lib="Soares breast 2NBHbst"  
/sex="Female"  
/dev\_stage="adult"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: breast; Vector: pMT73D (Pharmacia) with a  
modified polylinker; Site\_1: Not I; Site\_2: Eco RI; lstr  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
TGTACCAATCTGAAGTGGAGCGGCCCTTTTTTTTTTTTTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of a modified pT7m3 vector (Pharmacia).  
Library went through one round of normalization to a Col -  
230. Library constructed by Bento Soares and M.Fatima  
Bonaldo."

BASE COUNT 82 a 113 c 89 g 116 t 3 others  
ORIGIN

Query Match 65.6%; Score 308.8; DB 10; Length 403;  
Best Local Similarity 94.0%; Pred. No. 3.7e-69;  
Matches 375; Conservative 0; Mismatches 15; Indels 9; Gaps 5;

QY 8 gcaattttgtctgtccctgattcttcaggtcaccacattgaagcttcttagcagtcctg 67  
|||||  
Db 2 GCATTTTGTCTGTCTCCCTGATCTTCAGTCCACCACCATGAAGTTCTTAGCAGTCCTG 61

QY 68 gtactcttggaggtttccatcttcttggtctctgcccagaatccgacaacagctgtctca 127  
|||||  
Db 62 GNACTCTTGGGAGTTTCCATCTTCTTGGTCTCTGCCAGAAATCCGACACAGCTGTCTCA 121

QY 128 qctgaacagctatcagctactgctctgctgatgatgaagccctgatgctgaacact 187  
|||||  
Db 122 GCTGACACGATATCCAGTACTGGTCTCTGCTGATGATGAAGCCCTGATGCTGAACCACT 181

QY 188 gctgctgcaacaactgcgaccactgctgctcctaccactgcaaccacgctgcttacc 247  
|||||  
Db 182 GCTGCTGCAACCACTGGACCACCTGCTGCTCTTACCCTGCAACCACTGCTGCTTACC 241

QY 248 actgctgctaaagacattccagttttaccacaaatgggttggggtatccccgaatgg--ta 305  
|||||  
Db 242 ACTGCTCGTAAAGACATTCAGTTTACCCAAATGGGTGGGATCTCCGGAATGGGTAG'301

QY 306 gagtgtctccttag--atggaaatcagcttgagctctt--ctgcaattgggtcacaaacta- 360  
|||||  
Db 302 GACTGTGCTCCCTGAGGATGGGAATGAGCTTGAGCTTCTTGGCAATTTGGGTACAAANTAT 361

QY 361 ttcatgcttctctgtg--atttcatcacaactacttaacctt 397  
|||||  
Db 362 TTCAATGCTTCTCTGGATTTCATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 400

RESULT 11  
AZ694036 314 bp DNA linear GSS 18-DEC-2000  
LOCUS  
DEFINITION  
AST-2H8BG1003 Genetrap HL-60 Human Promyelocytic Leukemia Library.  
Homo sapiens genomic 5', DNA sequence.  
ACCESSION  
AZ694036  
VERSION  
AZ694036.1 GI:11878955

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

GSS.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 314)  
Henkel G., Liyanage, M., Pratt, E., Huang, D., Riley, M., Bernardino, A.,  
Durick, K. and Pollock, B.  
Exon-trap tags from a HL-60 GenomeScreen(TM) Library  
Unpublished (2000)  
Contact: Greg Henkel  
Gene Expression  
Aurora Biosciences Corp.  
11010 Torreyana Road, San Diego, CA 92121, USA  
Tel: 8584048436  
Fax: 8584046719  
Email: henkelg@aurorabio.com

Pools of cells were isolated from a GenomeScreen(TM) library. The  
library of cells was generated by retroviral integration of a gene  
tagging element consisting of: 1) A promoterless beta-lactamase  
preceded by a splice acceptor as a reporter for gene expression;  
2) A promoter driving neomycin resistance followed by a splice  
donor to trap downstream exons. 3' RACE from neomycin gene was  
performed using total RNA from isolated pools. Output was shotgun  
cloned in pAmp-1 and used to transform DH5-alpha competent  
bacteria. 5' ends of reported sequences were immediately preceded  
by splice donor from the trapping construct.  
Class: exon-trapped.

Location/Qualifiers  
1. 314  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Genetrap HL-60 Human Promyelocytic Leukemia  
Library"  
/tissue\_type="acute promyelocytic leukemia"  
/cell\_type="promyeloblast"  
/cell\_line="HL-60"  
/note="Organ: peripheral blood; Vector: pAmp-1; 3' RACE of  
total RNA from genetrap pools; shotgun clone in pAmp-1 and  
used to transform DH5-alpha competent bacteria."

BASE COUNT 76 a 88 c 56 g 94 t  
ORIGIN

Query Match 64.0%; Score 301.4; DB 12; Length 314;  
Best Local Similarity 99.4%; Pred. No. 2.9e-67;  
Matches 313; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 147 ctggctctgctgatgaagccctgctgctgaaccactgctgctgcaacaactgcga 206  
|||||  
Db 1 CTGGTCTCTGATGATGAAGCCCTGATGCTGAACCACTGCTGCTGAACCACTGCGA 60

QY 207 ccaactgctgctcctaccactgcaaccacccgctgcttctaccactgctgtaaacacattc 266  
|||||  
Db 61 CCACCTGCTCTCTACCCTGCAACCAACCGCTGCTTCTACCACCTGCTCGTAAAGACATTC 120

QY 267 cagttttaccacaaatgggttggggatctctcccaaatggttagagtgctgctcctgagatgaa 326  
|||||  
Db 121 CAGTTTATCCCAAAATGGGTGGGGATCTCCCGAATGGTAGAGTGTGCTCCTGAGATGAA 180

QY 327 tcagcttgagctctctgcaattgggtcacaactattcattcctgctgctgattctccaa 386  
|||||  
Db 181 TCAGCTTGAGTCTTCTGCNAAT--GGTCACAACATATTCATGCTTCTCTGATTTATCAACAA 239

QY 387 ctacttaccttgctacgatatcccccttttatctctaatcagtttattttcttccaaataa 446  
|||||  
Db 240 CTACTTACCTTGCCTACGATATCCCCCTTTATCTCTTAATCAGTTTATTTCTTTCAATAA 299

QY 447 aaaaataactatgagc 461  
|||||  
Db 300 AAAATAACTATGAGC 314

RESULT 12  
 BG202313/c  
 LOCUS  
 DEFINITION  
 RST21669 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
 ACCESSION  
 BG202313  
 VERSION  
 BG202313.1 GI:13724000  
 KEYWORDS  
 EST.  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 343)  
 Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
 Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,  
 Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith  
 ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher  
 ,J., Danzig,J. and Ducar,M.  
 Creation of genome-wide protein expression libraries using random  
 activation of gene expression  
 Nat. Biotechnol. 19 (5), 440-445 (2001)  
 21227151  
 CONTACT: Scott J. Cain  
 Athersys, Inc.  
 3201 Carnegie Ave, Cleveland, OH 44115, USA  
 Tel: 216 431 9900  
 Fax: 216 361 9596  
 Email: scain@athersys.com  
 High quality sequence stop: 343.  
 Location/Qualifiers  
 1..343  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Athersys RAGE Library"  
 /cell\_line="HT1080"  
 /note="See 'Creation of Genome-wide Protein Expression  
 Libraries using Random Activation of Gene Expression',  
 Nature Biotechnology, in press. Note that even though the  
 cell type indicated is HT1080, since a random activation  
 method was used, these sequence tags are not necessarily  
 expressed in HT1080 under normal circumstances."  
 99 a 63 c 97 g 80 t 4 others

FEATURES  
source

1..343  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Athersys RAGE Library"  
 /cell\_line="HT1080"  
 /note="See 'Creation of Genome-wide Protein Expression  
 Libraries using Random Activation of Gene Expression',  
 Nature Biotechnology, in press. Note that even though the  
 cell type indicated is HT1080, since a random activation  
 method was used, these sequence tags are not necessarily  
 expressed in HT1080 under normal circumstances."  
 99 a 63 c 97 g 80 t 4 others

BASE COUNT  
ORIGIN

Query Match 63.5%; Score 299.2; DB 10; Length 343;  
 Best Local Similarity 93.3%; Pred. No. 1.le-66;  
 Matches 321; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

OY 105 aqaatccgacacagctctccagctgacagtgatccagctactggtctctgctgatg 164  
 Db 343 AGAATCTGACACAGCTCTCCAGCTGACAGTATCCAGCTACTGGTCTCTCATGATG 284  
 OY 165 aagccctgagtgaaacacagctctgctgcaacaactgcgacactgctgctctacca 224  
 Db 283 AGCCCTGATGCTGAAACCACTGTTGGTGCAACANTGGCACCANTGCTCCGCCCC 224  
 OY 225 ctgcacacacagctgctcttaccactgctgctgaaagacattccagttttaccacaaatggg 284  
 Db 223 CTGCAACACCCGNTGCTTTTACCAGTCTGCTCTAAAGACATTTCCAGTTTACCCAAATGGG 164  
 OY 285 ttggggatctccgaaagttagagtggtgctccctgagatggaaatcagcttctctg 344  
 Db 163 TTGGGGATCTCCGAAATGGTGAAGTGTGTCCTGAGATGGAATCAGCTTGTCTCTGTC 104  
 OY 345 aattgggttcacaactattcatgctctctctggtatttcacaaactacttacctgctcag 404  
 Db 103 AATT-GGTTCACAACATTTATCATGCTCTCTGATGTTTATCTCCCTTACCTTCCCTTACG 45  
 OY 405 atatccctttatctctaactcagtttattttcttccaaataaaa 448  
 Db 44 ATATCCCTTTTATCTCTAATCAGTTTTTTTCTTCAATAAAAAA 1

RESULT 13  
 AW176044/c

LOCUS  
 DEFINITION  
 OVO-BT0103-270899-008-f12 BT0103 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION  
 AW176044  
 VERSION  
 AW176044.1 GI:6442081  
 KEYWORDS  
 EST.  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 313)  
 HCGP http://www.ludwig.org.br/ORESTES.  
 The FAPESP/LICR Human Cancer Genome Project  
 Unpublished (1999)  
 JOURNAL  
 COMMENT  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV0t2-QV0-BT0103-  
 270899-008-f12&t3-1999-08-27&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 7  
 High quality sequence stop: 313.  
 Location/Qualifiers  
 1..313  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="BT0103"  
 /dev\_stage="Adult"  
 /note="Organ: breast; Vector: puc18; Site\_1: SmaI; Site\_2:  
 SmaI; A mini-library was made by cloning products derived  
 from ORESTES PCR (U.S. Letters Patent application No. 196  
 ,716 - Ludwig Institute for Cancer Research) profiles  
 into the puc 18 vector. Reverse transcription of tissue  
 mRNA and cDNA amplification were performed under low  
 stringency conditions."

FEATURES  
source

1..313  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="BT0103"  
 /dev\_stage="Adult"  
 /note="Organ: breast; Vector: puc18; Site\_1: SmaI; Site\_2:  
 SmaI; A mini-library was made by cloning products derived  
 from ORESTES PCR (U.S. Letters Patent application No. 196  
 ,716 - Ludwig Institute for Cancer Research) profiles  
 into the puc 18 vector. Reverse transcription of tissue  
 mRNA and cDNA amplification were performed under low  
 stringency conditions."

BASE COUNT 93 a 57 c 88 g 75 t

## ORIGIN

Query Match 62.4%; Score 294; DB 9; Length 313;  
 Best Local Similarity 98.1%; Pred. No. 2.3e-65;  
 Matches 308; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

OY 136 gtatccagctactggtctgctgctgatgatgaagccctgatgctgaaacactgctgctgc 195  
 Db 313 GTATCCAGCTACTGCTGCTGCTGATGATGAAGCCCTGATGCTGAAACCACTGCTGCTGC 254  
 OY 196 aacaactgcacacactgctgctctcaccactgaacacacagctgctctcaccactgctgc 255  
 Db 253 AACCACTGCGACCACTGCTGCTCTACCACTGCAACACACCGCTGCTTACCACTGCTCG 194  
 OY 256 taaagacattccagttttaccacaaatgggttggggatctcccgaaatggtagagtgltcc 315  
 Db 193 TAAAGACATTCCAGTTTACCCAAATGGGTTGGGATCTCCCGAANTGGTAGAGTGGGTCC 134  
 OY 316 ctgagatggaaatcagcttgagcttctgcaattgggtgcacaaactattcatgcttctctg 375  
 Db 133 CTGAGATGGAATCAGCTTGAGTCTTCTGCAATT-GGTTCACAACATTTATCATGCTACCTGTG 75  
 OY 376 atttaccacactacttaccctgctcctacatattccctttatcttaataacagttatttt 435  
 Db 74 ATTTATCCCAACACTACTTACCTTGCCTACCATATCCCTTTATCTCTAATCAGTTTATTTT 15  
 OY 436 ctttcaataaaaaa 449  
 ||||||||| |



Db	14	CTTTCAANAATA 1
RESULT 14		
AA340069	294 bp	mRNA linear EST 21-APR-1997
LOCUS	EST45219	Fetal skin Homo sapiens cDNA 5' end similar to similar to
DEFINITION	Spot 2, submandibular gland protein, mRNA sequence.	
ACCESSION	AA340069	
VERSION	AA340069.1	GI:1992307
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	(bases 1 to 294)	
	Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult	
	,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White	
	,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,	
	Cline,T.R., Colton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald	
	,J.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,	
	Gnehm,C., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,	
	Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,	
	Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,	
	Phillips,C.A., Ryder,S.E., Scott,J.L., Saudak,D.M., Shirley,R.,	
	Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,	
	Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,	
	Dimke,D., Feng,D.-F., Ferrite,A., Fischer,C., Hastings,G.A., He,W.W.	
	, Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,	
	Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L.,	
	Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon	
	M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M., and	
	Venter,J.C.	
TITLE	Initial assessment of human gene diversity and expression patterns	
JOURNAL	based upon 83 million nucleotides of cDNA sequence	
MEDLINE	Nature 377 (6547 Suppl), 3-174 (1995)	
COMMENT	96026280	
	Other_ESTS: THCL17518	
Contact:	Kerlavage, AK	
Bioinformatics	The Institute for Genomic Research	
	9712 Medical Center Drive, Rockville, MD 20850 USA	
Tel:	3018699056	
Fax:	3018699423	
Email:	arkerlav@tigr.org	
For clone availability,	additional sequence and expression	
information related to this EST,	please check the TIGR Human Gene	
Index ( <a href="http://www.tigr.org/tdb/hgi/hgi.html">http://www.tigr.org/tdb/hgi/hgi.html</a> )		
Seq primer:	M3 Reverse.	
FEATURES	Location/Qualifiers	
source	1..294	
	/organism="Homo sapiens"	
	/db_xref="ATCC (inhost):141700"	
	/db_xref="taxon:9606"	
	/clone_lib="Fetal skin"	
	/tissue_type="epithelium"	
	/cell_type="epithelial cell"	
	/dev_stages="fetus"	
	/note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI	
	; Site_2: XhoI"	
BASE COUNT	60 a 91 c 58 g 82 t 3 others	
ORIGIN		
Query Match	61.4%; Score 289.4; DB 9; Length 294;	
Best Local Similarity	98.6%; Pred. No. 3.6e-64;	
Matches 290; Conservative	0; Mismatches 4; Indels 0; Gaps 0;	
QY	1	ctttgaagcattttgtctgctcccgatcttcagggtaccacacctgagtcttagc 60
Db	1	TTTTGAAGCATTTTNTCTGTCNCTCCCGATCTTCAGGTCCACCACCATGAAGTTCCTTAGC 60
QY	61	agtcctggactctctggaggtttccatcttctggtctctgcccaagaatccgacaacagc 120;





```
Db 22 ctttgaagcatctttgtctgtctccctgatctctcaggtcaccaccatgaagtcttagc 81
Qy 61 agtctctggtaactcttggaggtttccatcttctcgtctctgcccgaatccgacacagc 120
Db 82 agtctctggtaactcttggaggtttccatcttctcgtctctgcccgaatccgacacagc 141
Qy 121 tctccagctgacagctatccagctactggtctctgctgatgatgaagccctgatgctga 180
Db 142 tctccagctgacagctatccagctactggtctctgctgatgatgaagccctgatgctga 201
Qy 181 aaccactgctgtaacaaactcgaccactgctgctctaccactgaaccaccgctgc 240
Db 202 aaccactgctgtaacaaactcgaccactgctgctctaccactgaaccaccgctgc 261
Qy 241 tictaccactgctgtaacaaactcgaccactgctgctctaccactgaaccaccgctgc 300
Db 262 tictaccactgctgtaacaaactcgaccactgctgctctaccactgaaccaccgctgc 321
Qy 301 tggtagagtgctccctgagatggaatcagcttgagctcttgcgaattgggtcacaacta 360
Db 322 tggtagagtgctccctgagatggaatcagcttgagctcttgcgaattgggtcacaacta 380
Qy 361 tcatgcttccctgatttccaaactacttacccttgcttacgatatcccttttatctc 420
Db 381 tcatgcttccctgatttccaaactacttacccttgcttacgatatcccttttatctc 440
Qy 421 laalcagtttattttcttttcaataaaaaataaactatgagcgagctaaaca 470
Db 441 laalcagtttattttcttttcaataaaaaataaactatgagcgagctaaaca 490

RESULT 2
US-09-253-691-3/c
; Sequence 3, Application US/09253691
; Patent No. 6124100
; GENERAL INFORMATION:
; APPLICANT: Dong Kyu JIN
; TITLE OF INVENTION: Diagnostic Method and Kit for Neuropsychiatric Diseases
; FILE REFERENCE: 1942/36
; CURRENT APPLICATION NUMBER: US/09/253,691
; EARLIER APPLICATION NUMBER: 1999-02-22
; EARLIER FILING DATE: KR 98-6, 278
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Wordperfect 6.1/Windows
; SEQ ID NO 3
; LENGTH: 397
; TYPE: DNA
; ORGANISM: human
US-09-253-691-3

Query Match 14.6%; Score 68.6; DB 3; Length 397;
Best Local Similarity 61.5%; Pred. No. 1.9e-11;
Matches 110; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 75 tggaggtttccatcttcttggtctctgcccagaatccgacacagctgctccagctga 134
Db 397 TGGCCCTTTCACATGATGTGAACCTGTGCTGATAGGTCCCGCTGCTGCTGCTGCTG 338
Qy 135 cgtatccagctactggtctgctgatgatgaagccctgatgctgaaccactgctgctg 194
Db 337 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 278
Qy 195 caaacactgacacactgctctaccactgaaccaccgctgcttaccactgct 253
Db 277 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 219

RESULT 3
US-09-135-994-1/c
; Sequence 1, Application US/09135994A
```

```
; Patent No. 6280938
; GENERAL INFORMATION:
; APPLICANT: Ranum et al.
; TITLE OF INVENTION: SCA7 GENE AND METHODS OF USE
; FILE REFERENCE: University of Minnesota
; CURRENT APPLICATION NUMBER: US/09/135,994A
; EARLIER FILING DATE: 1998-08-18
; EARLIER APPLICATION NUMBER: 60/056,170
; EARLIER FILING DATE: 1997-08-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-135-994-1

Query Match 13.9%; Score 65.4; DB 4; Length 477;
Best Local Similarity 63.9%; Pred. No. 1.9e-10;
Matches 99; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 99 ctgcccagaatccgacacagctgctccagctgacacgtatccagctactggtctgctg 158
Db 300 CTGGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241
Qy 159 atgatgaagccctgatgctgaaccactgctgctgcaacaactgcgaccactgctgctc 218
Db 240 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181
Qy 219 ctaccactgcaaccaccgctgcttctaccactgct 253
Db 180 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 146

RESULT 4
US-08-469-802B-3/c
; Sequence 3, Application US/08469802B
; Patent No. 5741645
; GENERAL INFORMATION:
; APPLICANT: Orr, Harry T.
; APPLICANT: Ranum, Laura P.W.
; APPLICANT: Chung, Ming-yi.
; APPLICANT: Zoghbi, Huda Y.
; TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
; Patent No. 5741645
; TITLE OF INVENTION: Type 1 and Method for Diagnosis
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mueeting, Raasch, Gebhardt & Schwappach, P.A.
; STREET: 119 No. 5741645th Fourth Street, Suite 203
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,802B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueeting, Ann M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 110.00030101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1225
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
```



```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 196..1902
US-08-320-559-29

Query Match      13.8%; Score 64.8; DB 1: Length 3376;
Best Local Similarity 57.4%; Pred. No. 8.5e-10;
Matches 117; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 43 caccatgaagttcttagcagctctggtactcttggagagttccatcttcttggtctctgc 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 844 CACTTTATGTTCTCTGGAGCTTTTGAAGGTTTTTCCCTGTGCTCTTCATTAATTTGT 785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 103 ccagaaatccgacaaacagctgctccagctgacacgataccagctactgctgctgatga 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 784 GAGGCTTTGAAAACTGGTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGC 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 163 tgaagccctgagtgtaaacacactgctgctgcaacacactgacacactgctgctctac 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 724 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGC 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 223 cactgcaaccaccgctgctctac 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 664 TACTGCTGCTGCTGCTGCTGCTGC 641

RESULT 7
US-08-545-860D-29/c
; Sequence 29, Application US/08545860D
; Patent No. 6040140
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Region
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESS: No. 6040140ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545.860D
; FILING DATE: 07-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04496
; FILING DATE: 22-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10930
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/327,392
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/320,559
; FILING DATE: 11-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/062,443
```

```

; FILING DATE: 14-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,094
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,839
; FILING DATE: 27-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/805,093
; FILING DATE: 11-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca Esq., Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1262
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3376 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 196..1902
US-08-545-860D-29

Query Match      13.8%; Score 64.8; DB 3: Length 3376;
Best Local Similarity 57.4%; Pred. No. 8.5e-10;
Matches 117; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 43 caccatgaagttcttagcagctctggtactcttggagagttccatcttcttggtctctgc 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 844 CACTTTATGTTCTCTGGAGCTTTTGAAGGTTTTTCCCTGTGCTCTTCATTAATTTGT 785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 103 ccagaaatccgacaaacagctgctccagctgacacgataccagctactgctgctgatga 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 784 GAGGCTTTGAAAACTGGTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGC 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 163 tgaagccctgagtgtaaacacactgctgctgcaacacactgacacactgctgctctac 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 724 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGC 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 223 cactgcaaccaccgctgctctac 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 664 TACTGCTGCTGCTGCTGCTGCTGC 641

RESULT 8
PCT-US94-04496-29/c
; Sequence 29, Application PC/TUS9404496
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESS: Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```

: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/04496
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: DeLuca Esq., Mark
: REGISTRATION NUMBER: 33,229
: REFERENCE/DOCKET NUMBER: TJU-1242
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (215) 568-3100
: TELEFAX: (215) 568-3439
: INFORMATION FOR SEQ ID NO: 29:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3376 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 196..1902
:
PCT-US94-04496-29

```

[illegible]

```

RESULT      9
US-09-043-303-7/c
; Sequence 7, Application US/09043303
; Patent No. 6251589
; GENERAL INFORMATION:
; APPLICANT: TSUJII, Shoji
; APPLICANT: SANPEI, Kazujiro
; TITLE OF INVENTION: Method for Diagnosing Spinocerebellar Ataxia Type 2 and
; TITLE OF INVENTION: Primers Therefor
; FILE REFERENCE: 0760-0241P
; CURRENT APPLICATION NUMBER: US/09/043,303
; CURRENT FILING DATE: 1998-05-18
; EARLIER APPLICATION NUMBER: PCT/JP96/01999
; EARLIER FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 203
; TYPE: DNA
; ORGANISM: p-2093 plasmid
US-09-043-303-7

```

Query Match 13.7%; Score 64.6; DB 4; Length 203;  
Best Local Similarity 59.6%; Pred. NO. 2.1e-10;  
Matches 109; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

```

QY 89 ttctggctctgccagaataccgacaacacgctgctccagctgacagctatccagctact 14
  || | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 193 TTCCGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 134
QY 149 ggtctctgatgatgaagcccccctgatgctgaaacacacactgctgctgcaacaacclgcgacc 208
  || | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 74
QY 209 actgctctccactgcaaacaccgctgcttctaccactgctcgtaaagacattcc 268
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14
QY 269 gtt 271
  |||
Db 13 GTT 11

RESULT 10
US-08-469-802B-4/c
; Sequence 4, Application US/08469802B
; Patent No. 5741645
; GENERAL INFORMATION:
; APPLICANT: Orr, Harry T.
; APPLICANT: Rantum, Laura P.W.
; APPLICANT: Chung, Ming-yi
; APPLICANT: Zoghbi, Huda Y.
; TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
; Patent No. 5741645
; TITLE OF INVENTION: Type 1 and Method for Diagnosis
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Muetling, Raasch, Gebhardt & Schwappach, P.A.
; STREET: 119 No. 5741645th Fourth Street, Suite 203
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,802B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Muetling, Ann M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 110.00030101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1225
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 168 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-469-802B-4

Query Match 13.5%; Score 63.4; DB 1; Length 168;
Best Local Similarity 63.4%; Pred. No. 4.4e-10;
Matches 97; Conservative 0; Mismatches 56; Indels 0; Gaps

QY 101 gccagaatccgacaacagctgctccagctgacagctatccagctactgctgctgctgat 160
  ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163 GCCCGGAGCCCTGCTGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 104
QY 161 gatgaagcccttgatgctgaaaccactgctgctgcaacaactgcgaccactgctgctct 220
  ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```







```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-305-1217
: TELEFAX: 612-305-1228
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 195 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
US-08-267-803B-2

```

```

Query Match      13.5%; Score 63.4; DB 2; Length 195;
Best Local Similarity 63.4%; Pred No. 4.8e-10;
Matches 97; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 101 gccagaaatccgacacagctgctccagctgacacglatccagctactggtcctgctgat 160
Db 190 gccccggagccctgctgaggtgctgctgctgctgctgctgctgctgctgctgct 131
Qy 161 gatgaagccctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 220
Db 130 gctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 71
Qy 221 accactgcaacaccgctgctgctgctgctgctgctgctgctgctgctgctgct 253
Db 70 gctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 38

```

Search completed: October 5, 2002, 05:08:26  
Job time: 4146 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 04:00:45 ; Search time 1900.29 Seconds  
(without alignments)  
5186.787 Million cell updates/sec

Title: US-09-525-361A-23  
 perfect score: 471  
 Sequence: 1 ctttgaagcatttttctctg.....aactatagcgcagctaaccat 471

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_inv.*
```

## SUMMARIES

Result	Query	Score	Length	DB	ID	Description
No.	Match					

Query Match	95.28;	Score 448.4;	DB 6;	Length 497;
Best Local Similarity	98.58;	Pred. No. 1.6e-96;		

1	448.4	95.2	497	6	AR177347	Sequence	AR177347
2	448.4	95.2	578	6	AX282981	Sequence	AX282981
3	447	94.9	482	6	BD007431	Reagent a	BD007431
4	446.8	94.9	553	6	BD007432	Reagent a	BD007432
5	446.2	94.7	472	6	AX240679	Sequence	AX240679
6	446.2	94.7	473	6	AX3240680	Sequence	AX3240680
7	445.4	94.6	472	6	AX328498	Sequence	AX328498
8	442.2	93.9	488	6	AX067332	Sequence	AX067332
9	383.4	81.4	396	9	AF414087	Homo sapi	AF414087
10	300.8	63.9	308	6	AX3240676	Sequence	AX3240676
11	300.8	63.9	308	6	BD007429	Reagent a	BD007429
12	286.8	60.9	292	6	AX240677	Sequence	AX240677
13	269.8	57.3	273	6	AX328499	Sequence	AX328499
14	255.6	54.3	264	6	AX079476	Sequence	AX079476
15	255.6	54.3	266	6	AX079475	Sequence	AX079475
16	247.2	52.5	291	6	AX079474	Sequence	AX079474
17	202.8	43.1	229	6	BD007428	Reagent a	BD007428
18	187	39.7	133000	9	AX079842	Homo sapi	AX079842
19	187	39.7	188873	2	AC019223	Homo sapi	AC019223
20	187	39.7	22877	2	AC024091	Homo sapi	AC024091
21	186.4	39.6	201	6	AX240675	Sequence	AX240675
22	181.4	38.5	197	6	AX240678	Sequence	AX240678
23	181.4	38.5	197	6	BD007430	Reagent a	BD007430
24	80.8	17.2	469	10	RATSPOT1	Sequence	RATSPOT1
25	70.2	14.9	397	6	BD010479	Diagnosti	BD010479
26	69.6	14.8	1591	10	RNO6295	Rattus no	AC006295
27	69.6	14.8	68409	2	AC105975	Mus muscu	AC105975
28	69.6	14.8	185363	2	AC094466	Rattus no	AC094466
29	69.4	14.7	5727	5	AY045719	Larus arg	AY045719
30	68.8	14.6	666	8	AF413050S2	Zea mays	AF413051
31	68.6	14.6	46904	2	AC100911	Mus muscu	AC100911
32	67.8	14.4	98221	9	AC097839	Rattus no	AC097839
33	67.6	14.4	155780	9	HS336012	Rattus no	AL513498
34	67.6	14.4	175695	2	AL606844	Homo sapi	AL606844
35	67.2	14.3	180668	2	AC020857	Mus muscu	AC020857
36	66.8	14.2	110000	2	AC092202_1	Continuation (2 of	AC092202_1
37	66.8	14.2	173851	2	AC096966	Rattus no	AC096966
38	66.8	14.2	204152	2	AC092203	Mus muscu	AC092203
39	66.6	14.1	2960	3	AF057052	Mamestra	AF057052
40	66.4	14.1	624	3	AF139019	Cepaea ne	AF139019
41	66.4	14.1	236962	2	AL606742	Mus muscu	AL606742
42	66.2	14.1	22398	5	FR0271723	Fugu rubr	AJ271723
43	66.2	14.1	229380	2	AC079636	Mus muscu	AC079636
44	66	14.0	5719	10	NMCAT51	Mus muscu	AF051726
45	66	14.0	27613	9	AC002053	Cosmid cl	AC002053

## ALIGNMENTS

Matches 463; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 ctttgaagcattttgtgtgctccctgatcttcacagtcacaccaccatgaagtcttttagc 60

Db 22 CTTTGAAGCATTTTGTGTGCTGCCCTGTGATCTTCAGGTCAACACCATGAAGTCTTTAGC 81

Qy 61 agtccgtactcttgaggagtttccatcttcttctgtgtctctgcccagaaatccgacaacagc 120

Db 82 AGTCCTGGTACTCTTGGGAGTTTCCATCTTCTGGTCTCTGCCGAGAAATCCGACAACAGC 141

Qy 121 tgcctccagctgacacglatccagctactggtccctgctgagatgaagaagccctgatgctga 180

Db 142 TGCTCCAGCTGACACGATCCAGCTACTGCTGCTGCTGATGATGAAGCCCTGATGCTGA 201

Qy 181 aaccactgctgctgaacaaacgacacgactgctgctcctaccactgcaaccacccgctgc 240

Db 202 AACCACTGCTGCTGCAACCACTGCGACCACTGCTGCTTACCCTACCACTGCAACCACTGCG 261

Qy 241 tctaccactgctgctgaacaaacatccagctttaccacaaatgggttgggagatctcccgaa 300

Db 262 TTCTACCACCTGCTGCTGCAACCACTGCGACCACTGCTGCTTACCCTACCACTGCAACCACTGCG 321

Qy 301 lyytagagtgctccctgagatgaagaacagcttgagctctctcgaattggttcacaaacta 360

Db 322 TGGTAGAGTGCTGCTGCAACCACTGCGACCACTGCTGCTTACCCTACCACTGCAACCACTGCG 380

Qy 361 tcatgctcctgctgatttccacaaactacattacctgctcctacagatccccctttatctc 420

Db 381 TTCTATGCTTCTGCTGATTTCTCAACCACTGCTGCTTACCCTACCACTGCAACCACTGCG 440

Qy 421 taalcagcttatttcttccaaataaaaaataaactagcgagcgaactaaca 470

Db 441 TAATCAGTTTATTTCTTTCAAATAAAAAATAACTATGACGCAAAAAA 490

## RESULT 2

AX282981 AX282981 578 bp DNA linear PAT 03-NOV-2001

LOCUS Sequence 30 from Patent WO0175171.

DEFINITION AX282981

ACCESSION AX282981

VERSION AX282981.1 GI:16609909

KEYWORDS human.

SOURCE Homo sapiens

REFERENCE Eukaryota: Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE Houghton, R.L., Dillon, D.C., Moles, D.A., Xu, J., Zehentner, B. and

JOURNAL Methods, compositions and kits for the detection and monitoring of

PATENT: WO 0175171-A 30 11-OCT-2001;

CORIXA CORPORATION (US)

FEATURES Location/Qualifiers

source 1..578

/db\_xref="taxon:9606"

BASE COUNT 206 a 137 c 88 g 147 t

ORIGIN

Query Match 95.2%; Score 448.4; DB 6; Length 578;

Best Local Similarity 98.5%; Pred. No. 1.6e-96;

Matches 463; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 ctttgaagcattttgtgtgctccctgatcttcacagtcacaccaccatgaagtcttttagc 60

Db 17 CTTTGAAGCATTTTGTGTGCTGCCCTGTGATCTTCAGGTCAACACCATGAAGTCTTTAGC 76

Qy 61 agtccgtactcttgaggagtttccatcttcttctgtgtctctgcccagaaatccgacaacagc 120

Db 77 AGTCCTGGTACTCTTGGGAGTTTCCATCTTCTGGTCTCTGCCGAGAAATCCGACAACAGC 136

Qy 121 tgcctccagctgacacglatccagctactggtccctgctgagatgaagaagcccttgatgctga 180

Db 137 TGCTCCAGCTGACACGATCCAGCTACTGCTGCTGCTGATGAAGCCCTGATGCTGA 196

Qy 181 aaccactgctgctgaacaaacgacacgactgctcctaccactgcaaccacccgctgc 240

Db 197 AACCACTGCTGCTGCAACCACTGCGACCACTGCTGCTTACCCTACCACTGCAACCACTGCG 256

Qy 241 tctaccactgctgctgaacaaacatccagctttaccacaaatgggttgggagatctcccgaa 300

Db 257 TTCTACCACCTGCTGCTGCAACCACTGCGACCACTGCTGCTTACCCTACCACTGCAACCACTGCG 316

Qy 301 tgcctccagctgctgaacaaacgacacgactgctcctaccactgcaaccacccgctgc 360

Db 317 TGCTCCAGCTGCTGCTGCAACCACTGCGACCACTGCTGCTTACCCTACCACTGCAACCACTGCG 375

Qy 361 tcatgctcctgctgatttccacaaactacattacctgctcctacagatccccctttatctc 420

Db 376 TTCTATGCTTCTGCTGATTTCTCAACCACTGCTGCTTACCCTACCACTGCAACCACTGCG 435

Qy 421 taalcagcttatttcttccaaataaaaaataaactagcgagcgaactaaca 470

Db 436 TAATCAGTTTATTTCTTTCAAATAAAAAATAACTATGACGCAAAAAA 485

## RESULT 3

BD007431

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

OS

PN

PD

PF

PI

POLA

PI

R

STEVEN

PC

PC

CC

CC

FT

FT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 458; Conservative

94.9%; Score 447; DB 6; Length 482;

98.9%; Pred. No. 3.5e-96;

2; Mismatches 2; Indels 1; Gaps 1;



```
QY 1 ctttaagcatttttctgtctccctgactctccaggtcaccaccatgaagttcttagc 60
|||||
Db 9 CTTTGAAGCATTTTGTCTGTCTCCCTGATCTTTCATGTCACCAACATGAAGTTCTTAGC 68
|||||
QY 61 antctctggactcttgggagtttccatctttctgtctctgcccagaatccgacacagc 120
|||||
Db 69 AGTCTCGGTACTCTTTGGGAGTTTCCATCTTTCTGCTCTCTGCCAGAAATCCGACAAAGC 128
|||||
QY 121 tgcctcagctgacagctatccagctactggtcctctgctgatgatgaagccctgagtctga 180
|||||
Db 129 TGTCTCAGCTGACAGCTATCCAGCTACTTCCAGTCTTGGTCTCTGATGATGAAGCCCTCGATGCTGA 188
|||||
QY 181 aaccactgctgctgcaacaacacgacctgctgctcctaccactgcaaccacagctgc 240
|||||
Db 189 ACCACTGCTGTGCAACCACTGCGACCACTGCTCTCTTACCACTGCAACCACTGCTGCTC 248
|||||
QY 241 tctaccactgctcgtataaagacattccagtttttaccacaaatgggttgggagatcctccgaa 300
|||||
Db 249 TTCTACCACTGCTCGTAAAGACATTCAGAGTTTACCCAAATGGGTGGGATCTCCCGAA 308
|||||
QY 301 tggtagagtgctgctcccgagatggaatcagcttgagctctctcaattgggtggtcacaacta 360
|||||
Db 309 TGGTAGAGTGCTGCTCCGAGATGGAATCAGCTTGAGTCTTCTGCAATT-GGTCAACAATA 367
|||||
QY 361 ttcattgctccctgctgatttccatcccaactacttaccctgctcctaccactgcaaccacagctgc 420
|||||
Db 368 TTTCATGCTTCCCTGATTTTCATCCCACTACTTACCTTGCCCTACGATATCCCTTTATCTC 427
|||||
QY 421 taatcagttattttcttcttcaataaaaaataaactatgagcga 463
|||||
Db 428 TAATCAGTTTATTTCTTTTCAAAATAAAAAATAACTATGAGCAA 470
|||||

RESULT 6
AX240680 AX328498
LOCUS AX328498 473 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 6 from Patent WO0165262.
ACCESSION AX240680
VERSION AX240680.1 GI:15797663
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Billing-Medel, P.A., Cohen, M., Colpitts, T.L., Friedman, P.N.,
Gordon, J., Granados, E.N., Hodges, S.C., Klass, M.R., Kratochvil, J.D.,
Roberts-Rapp, L., Russell, J.C. and Stroupe, S.D.
JOURNAL Reagents and methods useful for detecting diseases of the breast
PATENT: WO 0165262-A 6 07-SEP-2001;
ABBOTT LABORATORIES (US)
FEATURES
Source Location/Qualifiers
1..473
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 108 a 134 c 86 g 145 t
ORIGIN
Query Match 94.7%; Score 446.2; DB 6; Length 473;
Best Local Similarity 99.1%; Pred. No. 5.4e-96;
Matches 459; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 ctttgaagcatttttctgtctccctgactctccaggtcaccaccatgaagttcttagc 60
|||||
Db 9 CTTTGAAGCATTTTGTCTGTCTCCCTGATCTTTCATGTCACCAACATGAAGTTCTTAGC 68
|||||
QY 61 agtctcgtactcttgggagtttccatctttctgtctctgcccagaatccgacacagc 120
|||||
Db 69 AGTCTCGGTACTCTTTGGGAGTTTCCATCTTTCTGCTCTCTGCCAGAAATCCGACAAAGC 128
|||||
QY 121 tgcctcagctgacagctatccagctactggtcctctgctgatgatgaagccctgagtctga 180
|||||
Db 129 TGTCTCAGCTGACAGCTATCCAGCTACTTCCAGTCTTGGTCTCTGATGATGAAGCCCTCGATGCTGA 188
|||||
QY 181 aaccactgctgctgcaacaacacgacctgctgctcctaccactgcaaccacagctgc 240
|||||
Db 189 ACCACTGCTGTGCAACCACTGCGACCACTGCTCTCTTACCACTGCAACCACTGCTGCTC 248
|||||
QY 241 tctaccactgctcgtataaagacattccagtttttaccacaaatgggttgggagatcctccgaa 300
|||||
Db 249 TTCTACCACTGCTCGTAAAGACATTCAGAGTTTACCCAAATGGGTGGGATCTCCCGAA 308
|||||
QY 301 tggtagagtgctgctcccgagatggaatcagcttgagctctctcaattgggtggtcacaacta 360
|||||
Db 309 TGGTAGAGTGCTGCTCCGAGATGGAATCAGCTTGAGTCTTCTGCAATT-GGTCAACAATA 367
|||||
QY 361 ttcattgctccctgctgatttccatcccaactacttaccctgctcctaccactgcaaccacagctgc 420
|||||
Db 368 TTTCATGCTTCCCTGATTTTCATCCCACTACTTACCTTGCCCTACGATATCCCTTTATCTC 427
|||||
QY 421 taatcagttattttcttcttcaataaaaaataaactatgagcga 463
|||||
Db 428 TAATCAGTTTATTTCTTTTCAAAATAAAAAATAACTATGAGCAA 470
|||||

RESULT 7
AX328498 AX328498
LOCUS AX328498 472 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 1 from Patent WO0135811.
ACCESSION AX328498
VERSION AX328498.1 GI:18098424
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Mack, D. and Gish, K.C.
Novel methods of diagnosing and determining prognosis of breast
cancer, compositions, and methods of screening for breast cancer
modulators
JOURNAL Patent: WO 0135811-A 1 25-MAY-2001;
EOS Biotechnology, Inc. (US)
FEATURES
Source Location/Qualifiers
1..472
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 109 a 132 c 88 g 143 t
ORIGIN
Query Match 94.6%; Score 445.4; DB 6; Length 472;
Best Local Similarity 99.2%; Pred. No. 8.4e-96;
Matches 469; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 1 ctttgaagcatttttctgtctccctgactctccaggtcaccaccatgaagttcttagc 60
|||||
Db 1 CTTTGAAGCATTTTGTCTGTCTCCCTGATCTTTCAGTCCACCACCATGAAGTTCTTAGC 60
|||||
QY 61 agtctcgtg--tactcttgggagtttccatctttctgtctctgcccagaatccgacacaa 118
|||||
Db 61 AGTCTGGTATACTCTTTGGGAGTTTCCATCTTCTTGCTCTCTGCCAGAAAT-GCACAACA 119
|||||
QY 119 gctgctccagctgacagctatccagctactggtcctgctgctgctgctgctgctgctgct 178
|||||
Db 120 GCTGCTCCAGCTGACAGCTATCCAGCTACTTGGTCTCTGCTGATGATGAAGCCCTGATGCT 179
|||||
QY 179 gaaaccactgctgctgcaacaacacgacctgctgctcctaccactgcaaccacagct 238
|||||
Db 180 GAAACCACTGCTGCTGCAACCACTGCGACCACTGCTGCTCTTACCACTGCAACCACTGCT 239
|||||
QY 239 gcttctaccactgctcgtataaagacattccagtttttaccacaaatgggttgggagatcctccg 298
|||||
Db 240 GCTTCTACCACTGCTCGTAAAGACATTCAGAGTTTACCCAAATGGGTGGGATCTCCCG 299
|||||
```

```
Db 129 TGCTCCACACTGACAGGTATCCAGCTACTGCTGCTGCTGATGATGAAGCCCTCTGATGCTGA 188
QY 181 aaccactgctgctgcaacaacacgacctgctgctcctaccactgcaaccacagctgc 240
|||||
Db 189 AACCACTGCTGCTCAACCACTGCGACCACTGCTGCTTACCACTGCAACCACTGCTGCTC 248
|||||
QY 241 tctaccactgctcgtataaagacattccagtttttaccacaaatgggttgggagatcctccgaa 300
|||||
Db 249 TTCTACCACTGCTCGTAAAGACATTCAGAGTTTACCCAAATGGGTGGGATCTCCCGAA 308
|||||
QY 301 tggtagagtgctgctcccgagatggaatcagcttgagctctctcaattgggtggtcacaacta 360
|||||
Db 309 TGGTAGAGTGCTGCTCCGAGATGGAATCAGCTTGAGTCTTCTGCAATT-GGTCAACAATA 367
|||||
QY 361 ttcattgctccctgctgatttccatcccaactacttaccctgctcctaccactgcaaccacagctgc 420
|||||
Db 368 TTTCATGCTTCCCTGATTTTCATCCCACTACTTACCTTGCCCTACGATATCCCTTTATCTC 427
|||||
QY 421 taatcagttattttcttcttcaataaaaaataaactatgagcga 463
|||||
Db 428 TAATCAGTTTATTTCTTTTCAAAATAAAAAATAACTATGAGCAA 470
|||||

RESULT 7
AX328498 AX328498
LOCUS AX328498 472 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 1 from Patent WO0135811.
ACCESSION AX328498
VERSION AX328498.1 GI:18098424
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Mack, D. and Gish, K.C.
Novel methods of diagnosing and determining prognosis of breast
cancer, compositions, and methods of screening for breast cancer
modulators
JOURNAL Patent: WO 0135811-A 1 25-MAY-2001;
EOS Biotechnology, Inc. (US)
FEATURES
Source Location/Qualifiers
1..472
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 109 a 132 c 88 g 143 t
ORIGIN
Query Match 94.6%; Score 445.4; DB 6; Length 472;
Best Local Similarity 99.2%; Pred. No. 8.4e-96;
Matches 469; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 1 ctttgaagcatttttctgtctccctgactctccaggtcaccaccatgaagttcttagc 60
|||||
Db 1 CTTTGAAGCATTTTGTCTGTCTCCCTGATCTTTCAGTCCACCACCATGAAGTTCTTAGC 60
|||||
QY 61 agtctcgtg--tactcttgggagtttccatctttctgtctctgcccagaatccgacacaa 118
|||||
Db 61 AGTCTGGTATACTCTTTGGGAGTTTCCATCTTCTTGCTCTCTGCCAGAAAT-GCACAACA 119
|||||
QY 119 gctgctccagctgacagctatccagctactggtcctgctgctgctgctgctgctgctgct 178
|||||
Db 120 GCTGCTCCAGCTGACAGCTATCCAGCTACTTGGTCTCTGCTGATGATGAAGCCCTGATGCT 179
|||||
QY 179 gaaaccactgctgctgcaacaacacgacctgctgctcctaccactgcaaccacagct 238
|||||
Db 180 GAAACCACTGCTGCTGCAACCACTGCGACCACTGCTGCTCTTACCACTGCAACCACTGCT 239
|||||
QY 239 gcttctaccactgctcgtataaagacattccagtttttaccacaaatgggttgggagatcctccg 298
|||||
Db 240 GCTTCTACCACTGCTCGTAAAGACATTCAGAGTTTACCCAAATGGGTGGGATCTCCCG 299
|||||
```



```

Qy 361 ttcattgcttctggtgatttcacccaactactactaccc 397
Db 360 TTCATGCTTCTGTGATTTTCATCCAACTACTTACCTT 396

RESULT 10
AX240676
LOCUS AX240676 308 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 2 from Patent WO0165262.
ACCESSION AX240676
VERSION AX240676.1 GI:15797659
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
REFERENCE
AUTHORS Billing-Medel,P.A., Cohen,M., Colpitts,T.L., Friedman,P.N.,
Gordon,J., Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D.,
Roberts-Rapp,L., Russell,J.C. and Stroupe,S.D.
TITLE Reagents and methods useful for detecting diseases of the breast
JOURNAL Patent: WO 0165262-A 2 07-SEP-2001;
ABBOTT LABORATORIES (US)
FEATURES
source
1..308
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 63 a 94 c 65 g 86 t
ORIGIN

Query Match 63.9%; Score 300.8; DB 6; Length 308;
Best Local Similarity 99.3%; Pred. No. 2.2e-61;
Matches 302; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ctttgaagcatttttctgtgctccctgatcttcaggtcaccaccatgaagttcttagc 60
Db 5 CTTTGAAGCATTTTCTGTGCTCCCTGATCTTCAGGTCCACCACCATGAAGTTCTTAGC 64

Qy 61 agtcctggactcttgaggagttccatcttctgctctcgtccagaaatccgacaacagc 120
Db 65 AGTCCTGGTACTCTTGGGAGTTTCCATCTTTCTGGTCTCTGCCAGAAATCCGACAACAGC 124

Qy 121 tgcctcagctgacacgtatcccaactactgctcgtcgtgatgaagccctgatgctga 180
Db 125 TGCCTCAGCTGACACGTATCCAGCTACTGGTCTGCTGATGATGAAGCCCTGATGCTGA 184

Qy 181 aaccactgctgcaacactgacacactgctcctcctaccactgcaaccaccgctgc 240
Db 185 AACCACTGCTGTGCAACCACTGCGACCACTGCTGCTCTACCACTGCAACCAACCGCTGC 244

Qy 241 tctaccactgctcgtaaagacattccagttttaccacaaatgggttggggatctcccgaa 300
Db 245 TTCTACCACTGCTCGTAAAGACATTCACAGTTTACCCAAATGGGTTGGGGATCTTCCGAA 304

Qy 301 tgggt 304
Db 305 TGGT 308

RESULT 11
LOCUS BD007429 308 bp DNA linear PAT 31-JAN-2002
DEFINITION Reagent and method useful in detecting mammary diseases.
ACCESSION BD007429
VERSION BD007429.1 GI:18635800
KEYWORDS JP 2001503980-A/2.
SOURCE
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 308)
AUTHORS Medel,P.A.B., Cohen,M., Colpitts,T.L., Friedman,P.N., Gordon,J.,

```

```

Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D., Rapp,L.R.,
Russell,J.C. and Stroupe,S.D.
Reagent and method useful in detecting mammary diseases
Patent: JP 2001503980-A 2 27-MAR-2001;
ABBOTT LABORATORIES
OS Unidentified
PN JP 2001503980-A/2
PD 27-MAR-2001
PF 31-OCT-1997 JP 1998520795
PR 31-OCT-1996 US 08/742067
PI PATRICIA A BILLING MEDEL,MAURICE COHEN,TRACY L COLPITTS, PI
POLA N FRIEDMAN,
PI JULIAN GORDON,EDWARD N GRANADOS,STEVEN C HODGES, PI MICHAEL
R KLASS,
PI JOHN D KRATOCHVIL,LISA ROBERTS RAPP,JOHN C RUSSELL, PI
STEVEN D STROUPE
PC C12N15/09,A61K39/395,A61K39/395,A61P35/00,C07K14/47,
C07K16/30,
PC C12N5/10,C12P21/02,C12Q1/68,C01N33/53,C12N15/00,C12N5/00, PC
A61K37/02
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..308
/organism='Unidentified'.
FEATURES
source
1..308
/organism='Unidentified'
/db_xref='taxon:32644'
BASE COUNT 63 a 94 c 65 g 86 t
ORIGIN

```

```

Query Match 63.9%; Score 300.8; DB 6; Length 308;
Best Local Similarity 99.3%; Pred. No. 2.2e-61;
Matches 302; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ctttgaagcatttttctgtgctccctgatcttcaggtcaccaccatgaagttcttagc 60
Db 5 CTTTGAAGCATTTTCTGTGCTCCCTGATCTTCAGGTCCACCACCATGAAGTTCTTAGC 64

Qy 61 agtcctggactcttgaggagttccatcttctgctctcgtccagaaatccgacaacagc 120
Db 65 AGTCCTGGTACTCTTGGGAGTTTCCATCTTTCTGGTCTCTGCCAGAAATCCGACAACAGC 124

Qy 121 tgcctcagctgacacgtatcccaactactgctcgtcgtgatgaagccctgatgctga 180
Db 125 TGCCTCAGCTGACACGTATCCAGCTACTGGTCTGCTGATGATGAAGCCCTGATGCTGA 184

Qy 181 aaccactgctgcaacactgacacactgctcctcctaccactgcaaccaccgctgc 240
Db 185 AACCACTGCTGTGCAACCACTGCGACCACTGCTGCTCTACCACTGCAACCAACCGCTGC 244

Qy 241 tctaccactgctcgtaaagacattccagttttaccacaaatgggttggggatctcccgaa 300
Db 245 TTCTACCACTGCTCGTAAAGACATTCACAGTTTACCCAAATGGGTTGGGGATCTTCCGAA 304

Qy 301 tgggt 304
Db 305 TGGT 308

RESULT 12
AX240677
LOCUS AX240677 292 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 3 from Patent WO0165262.
ACCESSION AX240677
VERSION AX240677.1 GI:15797660
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 292)

```



AUTHORS Billing-Medel, P.A., Cohen, M., Colpitts, T.L., Friedman, P.N.,  
Gordon, J., Grandos, E.N., Hodges, S.C., Klass, M.R., Kratochvil, J.D.,  
Roberts-Rapp, L., Russell, J.C., and Stroupe, S.D.  
TITLE Reagents and methods useful for detecting diseases of the breast.  
JOURNAL Patent: WO 0165262-A 3 07-SEP-2001;  
ABBOTT LABORATORIES (US)

FEATURES  
Source Location/Qualifiers

1..292  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="EST Clone 901429"

HASE COUNT 59 a 92 c 59 g 80 t 2 others  
ORIGIN

Query Match 60.9%; Score 286.8; DB 6; Length 292;  
Best Local Similarity 98.6%; Pred. No. 4.6e-58;  
Matches 288; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 8 gcatttttgcgtgctccctgattcttcagctcaccaccatgaagttcttagcagctctg 67  
|||||  
Db 1 GCATTTTGTCTGTCTCTCCCTGATCTTCATGTACACCATGAAGTTCTTAGCAGTCTG 60  
|||||  
Oy 68 gtactcttggagtttccatcttttgcgtctctgcccagaatccgacacagctgctcca 127  
|||||  
Db 61 GTACTCTTTGGGAGTTTCCATCTTTCTGCTCTCTGCCAGAAATCCGACAAACAGCTGCTCCA 120  
|||||  
Oy 128 gctgacagatccagctactggtccctgctgatgatgaagccctgactgctgaaccact 187  
|||||  
Db 121 CCGTACAGCTATCCAGCTACTGTGCTCTGCTGATGATGAAGCCCTGTGCTGAACCACT 180  
|||||  
Oy 188 gctgctgaacacactgacacactgctctctaccactgacacacagctgcttacc 247  
|||||  
Db 181 GCTGCTGAACCACTGCCACCACTGCTCTCTCTACCACTGCAACCACTGCTTTTATCC 240  
|||||  
Oy 248 actgctgtaagacatccagctgatttaccacaaatgggtgggactctccga 299  
|||||  
Db 241 ACTGCTGTAAGACATTNCAGTTTATCCCAATGGGTGGGATCTCCCA 292  
|||||

RESULT 13  
AX328499  
LOCUS  
DEFINITION Sequence 2 from Patent WO0135811.  
ACCESSION AX328499  
VERSION AX328499.1 GI:18098425  
KEYWORDS  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (sites)  
REFERENCE Mack, D. and Gish, K.C.  
AUTHORS Novel methods of diagnosing and determining prognosis of breast  
TITLE cancer, compositions, and methods of screening for breast cancer  
modulators  
JOURNAL Patent: WO 0135811-A 2 25-MAY-2001;  
EOS Biotechnology, Inc. (US)

FEATURES  
source Location/Qualifiers  
1..273  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 58 a 85 c 59 g 71 t  
ORIGIN

Query Match 57.3%; Score 269.8; DB 6; Length 273;  
Best Local Similarity 99.3%; Pred. No. 5.1e-54;  
Matches 271; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 47 atgaagtcttagcagctcctggtactcttggagtttccatcttctggtctgcccag 106  
|||||  
Db 1 ATGAAGTCTTAGCAGCTCCTGGTACTCTCTGGGAGTTTCCATCTTTCTGCTCTGCCCCAG 60  
|||||

Oy 107 aatccgacaaacagctgtctccagctgacacgtatccagctactgctcctgctgatgatgaa 166  
|||||  
Db 61 AATCCGACAAACAGCTGTCTCCAGCTGACACGTATCCAGCTACTGCTGCTGATGATCAA 120  
|||||  
Oy 167 gccctgatgctgaacacactgctgctgcaacactgacacactgctgctcctaccact 226  
|||||  
Db 121 GCCCTGATGCTGAACACCACTGCTGCTCAACCACTGCGACCACTGCTGCTCTTACCCT 180  
|||||  
Oy 227 gcaacacacgctgcttctaccacagctcgtgaaagacatccagtttaccacaaatggglt 286  
|||||  
Db 181 GCAACACCGCTGCTTCTTACCACCTGCTCGTAAACACATTCAGTTTACCCAAATGGGTT 240  
|||||  
Oy 287 ggggatctcccgaaatggtagtggtgctccctga 319  
|||||  
Db 241 GGGGATCTCCCGAATGGTAGAGTGTGCTCCCTGA 273  
|||||

RESULT 14  
AX079476  
LOCUS  
DEFINITION Sequence 220 from Patent WO0107611.  
ACCESSION AX079476  
VERSION AX079476.1 GI:13159036  
KEYWORDS  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 264)  
REFERENCE Baker, K.P., Goddard, A. and Wood, W.J.  
AUTHORS Human polypeptides and methods for the use thereof  
TITLE Patent: WO 0107611-A 220 01-FEB-2001;  
JOURNAL Genentech, Inc. (US)

FEATURES  
source Location/Qualifiers  
1..264  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 54 a 87 c 53 g 70 t  
ORIGIN

Query Match 54.3%; Score 255.6; DB 6; Length 264;  
Best Local Similarity 98.5%; Pred. No. 1.2e-50;  
Matches 258; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 ctltgaagcatttttgcgtgctccctgattctcagtcaccacatgaagttctttagc 60  
|||||  
Db 3 CTTTGAAGCATTTTTGTCTGTCTCCCTGATCTTTCAGTCAACCCCAATGAAGTTCTTAGC 62  
|||||  
Oy 61 agtctggtactcttggagtttccatcttctggtctctgcccagaatccgacacagc 120  
|||||  
Db 63 AGTCCTGCTACTCTTGGGAGTTTCCATCTTCTGCTCTGCCCAGAAATCCGACAAACAGC 122  
|||||  
Oy 121 tgcctcagctgacagctatccagctactggtcctgctgctgatgatgaagccctgactga 180  
|||||  
Db 123 TGCTCCAGCTGACACGTATCCAGCTACTGCTGCTGCTGATGATGAAGCCCTGATGCTGA 182  
|||||  
Oy 181 aaccactgctgctgcaacaaactgcgacactgctgctcctaccactgcaaccacgctgc 240  
|||||  
Db 183 AACCACTGCAACTGCAACCACTGCGACCACTGCTGCTCTTACCACCTGCAACCACTGCG 242  
|||||  
Oy 241 tctaccactgctcgtgaaagac 262  
|||||  
Db 243 TTCTACCACCTGCTCGTAAAGAC 264  
|||||

RESULT 15  
AX079475  
LOCUS  
DEFINITION Sequence 219 from Patent WO0107611.  
ACCESSION AX079475  
VERSION AX079475.1 GI:13159035



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 04:39:04 ; Search time 211.95 Seconds  
(without alignments)  
3815.365 Million cell updates/sec

Title: US-09-525-361A-23  
Perfect score: 471  
Sequence: 1 ctttgaacattttgtctg.....aactatgagcgagtaacat 471

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_032802.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	469.4	99.7	471	21	AAA54120
2	469.4	99.7	471	22	AAF85500
3	448.4	95.2	480	20	AAH84202
4	448.4	95.2	480	21	AAH79431
5	448.4	95.2	489	20	AAH00704
6	448.4	95.2	489	20	AAH00644
7	448.4	95.2	497	20	AAZ40783
8	448.4	95.2	578	21	AAH79470
9	448.4	95.2	578	22	AAH67224

10	447	94.9	482	19	AAV31992	BS106 polynucleoti
11	447	94.9	482	20	AAH55581	BS106 consensus po
12	446.8	94.9	553	19	AAV31993	BS106 polynucleoti
13	446.8	94.9	553	20	AAH55582	BS106 consensus po
14	446.4	94.8	544	21	AAZ91765	Human breast speci
15	446.2	94.7	472	22	AAH18673	Human BS106 gene-s
16	446.2	94.7	472	22	AAH18674	Human BS106 gene-s
17	443	94.1	471	21	AAZ65099	Membrane-bound pro
18	443	94.1	471	22	AAH44245	Human PRO1160 (UNQ
19	443	94.1	471	22	AAH44245	Human anglogenesis
20	442.2	93.9	488	22	AAH44870	Human breast cancer
21	300.8	63.9	308	19	AAV31990	BS106 polynucleoti
22	300.8	63.9	308	20	AAH55579	BS106 clone 893988
23	300.8	63.9	308	22	AAH18670	Human BS106 gene-s
24	298.8	63.4	501	23	AAH87721	DNA encoding novel
25	286.8	60.9	292	22	AAH18671	Human BS106 gene-s
26	271.4	57.6	273	21	AAH54121	Breast cancer prot
27	255.6	54.3	264	22	AAH93399	CDNA encoding SRT
28	255.6	54.3	266	22	AAH93398	CDNA encoding SRT
29	247.2	52.5	291	22	AAH93397	BS106 polynucleoti
30	202.8	43.1	229	19	AAH31989	BS106 clone 166288
31	202.8	43.1	229	20	AAH55578	Human BS106 gene-s
32	186.4	39.6	201	22	AAH18669	BS106 polynucleoti
33	181.4	38.5	197	19	AAV31991	BS106 clone 120981
34	181.4	38.5	197	20	AAH55580	Human BS106 gene-s
35	181.4	38.5	197	22	AAH18672	DNA encoding novel
36	94	20.0	660	23	AAH87720	Spinocerebellar at
37	70.2	14.9	397	20	AAH89891	Human SCA7 genomic
38	65.4	13.9	477	21	AAZ44307	Spinocerebellar at
39	65	13.8	234	16	AAH84832	DNA encoding novel
40	64.8	13.8	1326	23	AAH83104	AF-9 cDNA. Homo s
41	64.8	13.8	3376	16	AAH75166	Glutamine rich reg
42	64.6	13.7	203	19	AAV30271	SCA2 gene CAG repe
43	64.6	13.7	203	19	AAH17226	Exons E, C and A o
44	64.6	13.7	1037	21	AAH55242	Exons D, C, B and
45	64.6	13.7	1472	21	AAH59241	

## ALIGNMENTS

RESULT 1  
AAH54120  
ID AAA54120 standard; DNA; 471 BP.  
XX  
AC AAA54120;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Breast cancer protein BCH1 coding sequence.  
XX  
KW Breast cancer; diagnosis; prognosis; detection; screening;  
KW antibody; oestrogen receptor; anti-oestrogen; immune response;  
KW lymph node; metastases; tumour; BCR3; BCQ5; BCH1; BCN1; BCN2;  
KW BCN5; BCQ5; BCX2; BCX3; BCA2; BCR2; BCJ7; BCY3; human; ds.  
XX  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 47..319  
FT */\*tag= a*  
FT */product= BCH1 protein*  
XX  
PN WO200055629-A2.  
PD 21-SEP-2000.  
XX  
PP 15-MAR-2000; 2000WO-US06952.  
XX  
PR 15-MAR-1999; 99US-0268865.  
PR 12-NOV-1999; 99US-0439878.  
PR 12-NOV-1999; 99US-0440370.  
PR 15-NOV-1999; 99US-0440493.

PR	16-NOV-1999;	99US-0440676.
PR	16-NOV-1999;	99US-0440677.
PR	29-NOV-1999;	99US-0450810.
PR	02-DEC-1999;	99US-0453137.
PR	08-MAR-2000;	2000US-0453137.
XX		
PA	(EOSB-) EOS BIOTECHNOLOGY INC.	
XX		
PI	Mark D, Gish KC;	
XX		
DR	WPI; 2000-638216/61.	
DR	P-PSDH; AAB00184.	
XX		
PT	Screening drug candidates for their ability to modulate breast cancer	
PT	by contacting the drug to a cell expressing an expression profile gene	
XX	and determining modulation of expression of the gene	
US	Disclosure; Fig 32; 258pp; English.	
XX		
CC	New methods for screening drug candidates are described which	
CC	comprise adding a drug candidate to a cell that expresses a protein	
CC	selected from BClH1, BClA2, BClJ7, BClN1, BClN5, BClO2, BClQ5, BClR2, BClX2	
CC	and BClY3 or their fragments and determining the effect of the drug	
CC	on the expression of those proteins. Antibodies to breast cancer	
CC	genes (specifically BClH1 or its fragment (BClH1p1 or BClH1p2)) are	
CC	useful for inhibiting and treating breast cancer in individuals who	
CC	are non-responsive to anti-oestrogen and positive for oestrogen	
CC	receptor. Compositions comprising BClH1 or a nucleic acid encoding	
CC	BClH1 are useful for eliciting an immune response in an individual.	
CC	The antibodies are also useful for the diagnosis and prognosis of	
CC	breast cancer and for screening compositions which modulate the	
CC	breast cancer phenotype. The method allows rapid and simple	
CC	detection of lymph node metastases.	
XX		
XX	Sequence 471 BP; 108 A; 133 C; 88 G; 142 T; 0 other;	

Query Match	99.7%	Score 469.4	DB 21	Length 471
Rest Local Similarity	99.8%	Pred. No. 3.4e-119		
Matches 470	Conservative 0	Mismatches 1	Indels 0	Gaps 0
Oy	1	cttgaagcattttgtctgtgctccctgatctctcaggtcacccaccatgaagttttttagc	60	
Db	1	cttgaagcattttgtctgtgctccctgatctctcaggtcacccaccatgaagttttttagc	60	
Oy	61	agtcttggtactcttggaggtttccattcttctggtctctgccagaatccgacaacagc	120	
Db	61	agtcttggtactcttggaggtttccattcttctggtctctgccagaatccgacaacagc	120	
Oy	121	tgtccagctgacacgtatccagctactggtctctgctgatgatgaagccctgatgctga	180	
Db	121	tgtccagctgacacgtatccagctactggtctctgctgatgatgaagccctgatgctga	180	
Oy	181	aaccactgctctgacaacaactgcgacactgctgctctctaccactgcacacacgcgtgc	240	
Db	181	aaccactgctctgacaacaactgcgacactgctgctctctaccactgcacacacgcgtgc	240	
Oy	241	tctaccactgctctgataagacattccagttttacccaaatgggtggggaatcccga	300	
Db	241	tctaccactgctctgataagacattccagttttacccaaatgggtggggaatcccga	300	
Oy	301	tgttagagtggtcccttgagatgggaatcagcttgatctcttcgaaatgggtcacaaacta	360	
Db	301	tgttagagtggtcccttgagatgggaatcagcttgatctcttcgaaatgggtcacaaacta	360	
Oy	361	tcatgcttctcttgatcttcatacaactacttaccctggctacgatatccccctttatctc	420	
Db	361	tcatgcttctcttgatcttcatacaactacttaccctggctacgatatccccctttatctc	420	
Oy	421	taatcagtttatctttcttcaataaaaaataactatgacggagctaaact	471	
Db	421	taatcagtttatctttcttcaataaaaaataactatgacggagctaaact	471	

RESULT	2	
AAF85500		
ID	AAF85500 standard; cDNA: 471 BP.	
XX		
AC	AAF85500;	
XX		
DT	23-JUL-2001 (first entry)	
XX		
DE	Nucleotide sequence of a human breast cancer protein designated BCH1.	
XX		
KW	Breast cancer protein; BCH1; breast cancer; chromosome 12; 12q12;	
KW	anti-oestrogen therapy; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	47..319
FT		/*tag= a
FT		/product= "breast cancer protein BCH1"
FT	sig_peptide	47..103
FT		/*tag= b
XX		
PN	WO200135811-A2.	
XX		
PD	25-MAY-2001.	
XX		
PF	16-NOV-2000; 2000WO-US31736.	
XX		
PR	16-NOV-1999; 99US-0440676.	
XX		
PA	(EOSB-) EOS BIOTECHNOLOGY INC.	
XX		
PI	Mack D, Gish KC;	
XX		
DR	WPI: 2001-343670/36.	
DR	P-PSDB; AAB83832.	
XX		
PT	BCH1 genes and protein useful for the diagnosis and treatment of breast	
PT	cancer -	
XX		
PS	Disclosure; Fig 1; 73pp; English.	
XX		
CC	The present sequence encodes a human breast cancer protein, designated	
CC	BCH1. BCH1 is upregulated in breast cancer tissue, and is found on	
CC	chromosome 12, cytoband 12q12. BCH1 can be used as an indicator of	
CC	breast cancer, for determining non-responsiveness to anti-oestrogen	
CC	therapy and for treating breast cancer. BCH1 nucleic acid is useful for	
CC	diagnosing breast cancer, and BCH1 is useful for determining the	
CC	prognosis of breast cancer and for determining whether an individual with	
CC	breast cancer will be responsive to anti-oestrogen therapy (where the	
CC	patient is positive for oestrogen receptor), where high levels indicate	
CC	poor prognosis and non-responsiveness, respectively. BCH1 is also useful	
CC	for screening for candidate drugs and bioactive agents. Inhibitors,	
CC	antibodies and antisense sequences of BCH1 are useful for treating	
CC	breast cancer. Antibodies to BCH1 are useful for localizing a therapeutic	
CC	moiety (e.g. cytotoxic agent or radioisotope) to breast cancer tissue,	
CC	and for treating breast cancer.	
XX		
SQ	Sequence 471 BP; 108 A; 133 C; 88 G; 142 T; 0 other;	

[illegible]



















```
PF 28-FEB-2001: 2001WO-US06516.
XX
PR 29-FEB-2000: 2000US-0516444.
XX
XX (ABBO ) ABBOTT LAB.
XX
PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX
XX WPI: 2001-596773/67.
XX
XX A gene or its fragment which codes for a BS106 polypeptide, useful for
XX the detection of a breast disease such as breast cancer
XX
XX Example 1: Fig 1: 158pp; English.
XX
XX The invention relates to human BS106-specific polypeptides and
XX polynucleotides. The BS106 polypeptides and antibodies are useful
XX for detecting, diagnosing, staging, monitoring, prognosticating,
XX preventing, treating or determining the predisposition of an
XX individual to diseases and conditions of the breast such as breast
XX cancer. They are also useful in the treatment of tumours or
XX metastases. Polynucleotides of the invention are useful in drug
XX screening and gene therapy. The present sequence is human BS106
XX gene specific full length cDNA, expressed sequence tag (EST) clone
XX 16628851nh.
XX
SQ Sequence 472 BP; 108 A; 134 C; 86 G; 144 T; 0 other;
```

Query Match 94.7%; Score 446.2; DB 22; Length 472;  
Best Local Similarity 99.1%; Pred. No. 8e-113;  
Matches 459; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```
OY 1 ctttgaagcatttttctgtgtcctcctgatcttcaggtcaccaccatgaagtctcttagc 60
Db 9 ctttgaagcatttttctgtgtcctcctgatcttcaggtcaccaccatgaagtctcttagc 68
OY 61 agtcctgttactcttggagggttccatcttctgtctgtccagaaatccgacaacagc 120
Db 69 agtcctgttactcttggagggttccatcttctgtctgtccagaaatccgacaacagc 128
OY 121 tgcctcagctgacagctatccagctactggtcctctgatgatgaagccctctgactga 180
Db 129 tgcctcagctgacagctatccagctactggtcctctgatgatgaagccctctgactga 188
OY 181 aaccactgctgtgcaacaactgcgaccactgctcctcctaccactgcaaccacgcctgc 240
Db 189 aaccactgctgtgcaaccactgcgaccactgctcctcctaccactgcaaccacgcctgc 248
OY 241 tctaccactgctgtaagacatctccagtttaccctcccaaatgggtggggtctcccgaa 300
Db 249 tctaccactgctgtaagacatctccagtttaccctcccaaatgggtggggtctcccgaa 308
OY 301 tggtagagtgtcctcctgagatgggaatcagcttgagttctctgcaattgggtcacaaacta 360
Db 309 tggtagagtgtcctcctgagatgggaatcagcttgagttctctgcaatt-ggtcacaaacta 367
OY 361 ttcatgtcttctgtgatttccatcccaactacttacccttgcctacgataccctttatctc 420
Db 368 ttcatgtcttctgtgatttccatcccaactacttacccttgcctacgataccctttatctc 427
OY 421 taatcagttattttcttcaataaaaaataaataatgatgagca 463
Db 428 taatcagttattttcttcaataaaaaataaataatgatgagcaad 470
```

Search completed: October 5, 2002, 05:44:20  
Job time: 3916 sec